The EMBO Journal vol.14 no.18 pp.4387–4394, 1995

The three-dimensional structure of the native ternary complex of bovine pancreatic procarboxypeptidase A with proproteinase E and chymotrypsinogen C

F.X.Gomis-Rüth[1], M.Gómez, W.Bode[2], R.Huber[2] and F.X.Avilés[3]

Institut de Biologia Fonamental i Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain and [2]Abteilung für Strukturforschung, Max-Planck-Institut für Biochemie, 82152 Martinsried, Germany
[1]Present address: Abteilung für Strukturforschung, Max-Planck-Institut für Biochemie, 82152 Martinsried, Germany

[3]Corresponding author

Dedicated to Hans Neurath, who played a primary role in the characterization of oligomeric forms of procarboxypeptidase A 40 years ago.

The metalloexozymogen procarboxypeptidase A is mainly secreted in ruminants as a ternary complex with zymogens of two serine endoproteinases, chymotrypsinogen C and proproteinase E. The bovine complex has been crystallized, and its molecular structure analysed and refined at 2.6 Å resolution to an *R* factor of 0.198. In this heterotrimer, the activation segment of procarboxypeptidase A essentially clamps the other two subunits, which shield the activation sites of the former from tryptic attack. In contrast, the propeptides of both serine proproteinases are freely accessible to trypsin. This arrangement explains the sequential and delayed activation of the constituent zymogens. Procarboxypeptidase A is virtually identical to the homologous monomeric porcine form. Chymotrypsinogen C displays structural features characteristic for chymotrypsins as well as elastases, except for its activation domain; similar to bovine chymotrypsinogen A, its binding site is not properly formed, while its surface located activation segment is disordered. The proproteinase E structure is fully ordered and strikingly similar to active porcine elastase; its specificity pocket is occluded, while the activation segment is fixed to the molecular surface. This first structure of a native zymogen from the proteinase E/elastase family does not fundamentally differ from the serine proproteinases known so far.

Keywords: chymotrypsinogen C/procarboxypeptidase A/proproteinase E/X-ray crystal structure/zymogen activation

Introduction

Proenzymes from the exocrine pancreas constitute useful models for proteolytic processing and enzyme regulation (Neurath, 1989) and many of them have been characterized in depth at both the structural and functional levels (trypsinogen, chymotrypsinogen A, procarboxypeptidases A and B, etc.). Exceptions are chymotrypsinogen C and proproteinase E, which are associated with procarboxypeptidase A in a ternary complex in ruminants (Brown *et al.*, 1961; Puigserver *et al.*, 1986, Avilés *et al.*, 1993). Conformational studies of the isolated subunits or of the whole ternary complex have been limited in the past because of the strong association between subunits in the complex, because of its large size (100 kDa) and because of its resistance to crystallization, a fact probably related to the rapid autolysis of the proproteinase E subunit (Pascual *et al.*, 1990).

The central subunit of the bovine pancreatic ternary complex (TC) is procarboxypeptidase A (PCPA), a 403-residue zymogen. The three-dimensional structure of its enzyme moiety, carboxypeptidase A (CPA), as well as its enzymatic and physico-chemical properties, have been described in great detail (see Avilés *et al.*, 1993 for a review). CPA is a 309-residue zinc-dependent metalloproteinase that preferentially cleaves peptide substrates with aliphatic or aromatic residues from the C-terminus. The conformation of the long activation segment (AS) of PCPA is assumed to be similar to that found in the crystal structures of the monomeric A and B zymogen homologues from porcine pancreas recently described (Coll *et al.*, 1991; Guasch *et al.*, 1992). In these species the AS almost completely shields the active-site depression. The multistep tryptic activation process (Vendrell *et al.*, 1990) liberates the mature enzyme without further significant conformational changes in the readily preformed active site. Besides its occurrence as a ternary complex with chymotrypsinogen C and proproteinase E, pancreatic PCPA is also found in a binary complex with either digestive serine endoproproteinases or as a monomeric form in non-ruminants (see Avilés *et al.*, 1993 for a review).

Bovine chymotrypsinogen C (CTGC) subunit, a single-chain zymogen of 251 residues, is the precursor of a chymotrypsin-like serine endopeptidase that cleaves preferentially after aliphatic residues, releasing peptides that are substrates of CPA. The amino-acid sequence of CTGC exhibits a stronger homology with proelastase than with chymotrypsinogen A (CTGA; Hartley, 1964), but shares with the latter proenzyme the disulfide-bond pattern. Upon tryptic activation, the 13-residue activation peptide is cleaved, but not liberated, remaining connected by a disulfide bridge to the enzyme moiety.

Bovine proproteinase E (BPE) subunit is the 253-residue precursor of an enzyme that belongs to the elastase/proteinase E family of serine proteinases that cleave polypeptides preferentially C-terminal of Ala, Val, Ser and Thr (Kobayashi *et al.*, 1981a,b). The primary structure of BPE was derived from an N-terminally truncated form of 240 residues (Venot *et al.*, 1986) and from the sequence of the intact N-terminus (Pascual *et al.*, 1990). It is highly homologous (86% sequence identity) to human pancreatic proproteinase E (Shen *et al.*, 1987) and moderately homo-

logous (~55% identity) to pancreatic elastases (Venot *et al.*, 1986). BPE is activated by trypsin by cleaving off the 11 amino acid residue activation peptide (Pascual *et al.*, 1990). Ruminant BPE can autolyse during certain isolation procedures, giving rise to the above mentioned N-terminal truncated form, called subunit III, which lacks the first two residues of the mature enzyme (Pascual *et al.*, 1990) and shows a strongly diminished activity (Wicker and Puigserver, 1981). This subunit has been assumed for many years to be an original component of the ruminant ternary complexes (Brown *et al.*, 1961; Kerfelec *et al.*, 1986); now it seems clear, however, that it is an isolation artefact (Pascual *et al.*, 1990; Avilés *et al.*, 1993). The three-dimensional structure of the monomeric subunit III from the bovine digestive system has recently been solved (Pignol *et al.*, 1994) and was considered to be a homology model for the proteinase E zymogens.

The *in vitro* activation of PCPA in the bovine TC is much slower than that observed for monomeric forms of PCPA in other species (Freisheim *et al.*, 1967; Puigserver *et al.*, 1986; Vendrell *et al.*, 1990) and has proven to be significantly enhanced by the presence of Ca^{2+} ions (Uren and Neurath, 1972; Puigserver and Desnuelle, 1977). It has been shown that, on tryptic attack, the chymotrypsin and proteinase E activities are released prior to the CPA activity without concomitant dissociation of the complex (Uren and Neurath, 1972). Various hypotheses about the biological function of the heterotrimer have been put forward, such as its assistance in timing, potentiation and targeting of the proteolytic activities, modulation of the activity of the subunits, proper coordination of the appearance of the proteolytic activities in the duodenum, and the protection of PCPA against inactivation by the acidic gastric juice (Uren and Neurath, 1972; Puigserver and Desnuelle, 1977; Michon *et al.*, 1991a,b; Avilés *et al.*, 1993).

In this publication we report the spatial structure of the intact 100 kDa bovine TC and the essential conformational properties of the three constituent zymogens. It is the only digestive complex of proprote(in)ases, for which the detailed three-dimensional structure has been solved so far. The structure of this complex explains functional data obtained in solution. With BPE, we report the first three-dimensional model of an activatable member of the proelastase/proproteinase E family.

Results and discussion

Arrangement of the TC in the crystal

Figure 1A displays the packing of the 3×9 subunits present in the hexagonal unit cell (with one heterotrimer per asymmetric unit). The TC is unequivocally assigned, as shown by the colour coding in Figure 1A, due to the dominating number of intracomplex versus packing contacts. The central PCPA molecule makes a much bigger number of atom–atom contacts with the adjacent BPE and CTGC molecules of a given TC, comparable in number and quality with typical protein–protein interfaces, than with molecules of neighbouring heterotrimers that resemble those usually observed in crystal contacts. Both serine proproteinases in the TC do not directly touch one another, but flank the central PCPA (see Figure 1B).

The structure of PCPA and comparison with the porcine homologous (monomeric) form

Bovine PCPA subunit, as the recently published porcine homologous monomeric structure (Guasch *et al.*, 1992) and that of porcine monomeric procarboxypeptidase B (PCPB; Coll *et al.*, 1991), consists of two distinct domains, the larger 309-residue enzyme moiety and the smaller 94-residue AS (see Figure 1B). The enzyme moiety is an α/β protein, with a central mixed β-sheet flanked on both sides by several α-helices, and with the active-site residues located close to the C-terminus of the central parallel β-strands. The AS consists of a globular portion, residues Lys4A–Glu80A (Coll *et al.*, 1991; Guasch *et al.*, 1992; nomenclature according to porcine PCPA, see PDB access code 1pca) and an adjacent connecting region (Asp81A–Arg99A). The globular part is a two-layer structure formed by a four-stranded antiparallel β-pleated sheet, covered on the surface by two helices. It exhibits a curled loop, which is directed towards the catalytic zinc ion and blocks access of larger substrates to the active site. The segment connecting the AS to the enzyme moiety consists of a three-turn α-helix and an adjacent large activation loop. This activation loop is well defined in this bovine form, in contrast to the monomeric porcine homologous form, probably due to a stabilization by the CTGC in the TC (see below). The potential primary tryptic cleavage site around the scissile peptide bond, Arg99A–Ala1, is slightly exposed, but integrated in an overall extended segment. The second cleavage site, at Arg74A–Tyr75A, is an integral part of the edge strand of the AS β-sheet.

The bovine proenzyme is very similar to monomeric porcine PCPA (Guasch *et al.*, 1992), but the C-terminus is longer by one residue, Tyr309. Superposition shows a slight relative rotation between the AS and the enzyme domain.

The structure of CTGC and comparison with related (pro)enzymes

The three-dimensional structure of the CTGC subunit is the first solved for a C form of the chymotrypsinogen family; it displays the characteristic architecture of chymotrypsin-like serine endopeptidases (Bode and Huber, 1986). It is composed of two interacting β-barrel-like domains, each consisting of a six-stranded, twisted β-pleated sheet, covered by a short 'intermediate' helical segment (Asp864–Ser869; for the numbering convention of CTGC residues, see Materials and methods), the extended C-terminal helix (Ile935–Leu944), and several surface-located loop segments. Some of these are arranged around the active-site triad, consisting of Ser895, His757 and Asp802, and shape the active-site groove (see Figure 2A).

Four peptide segments, which after activation cleavage form part of the specificity pocket or are in its vicinity, are organized in a zymogen-like manner. In the active enzymes, this pocket mainly consists of the 'entrance frame' (comprising the residues equivalent to Val913–Cys920), the 'base' Ser889–Ser895, the connecting disulfide bridge Cys891–Cys920, and the 'back'-forming segment Thr926–Phe928. Similar to CTGA, but in strong contrast to the active enzymes, Ala890–Asp894 is folded inwards in CTGC (see Figure 2A), in this way blocking the pocket for substrates. Simultaneously, the side-chain of Asp894 is directed away from the molecule, forming,

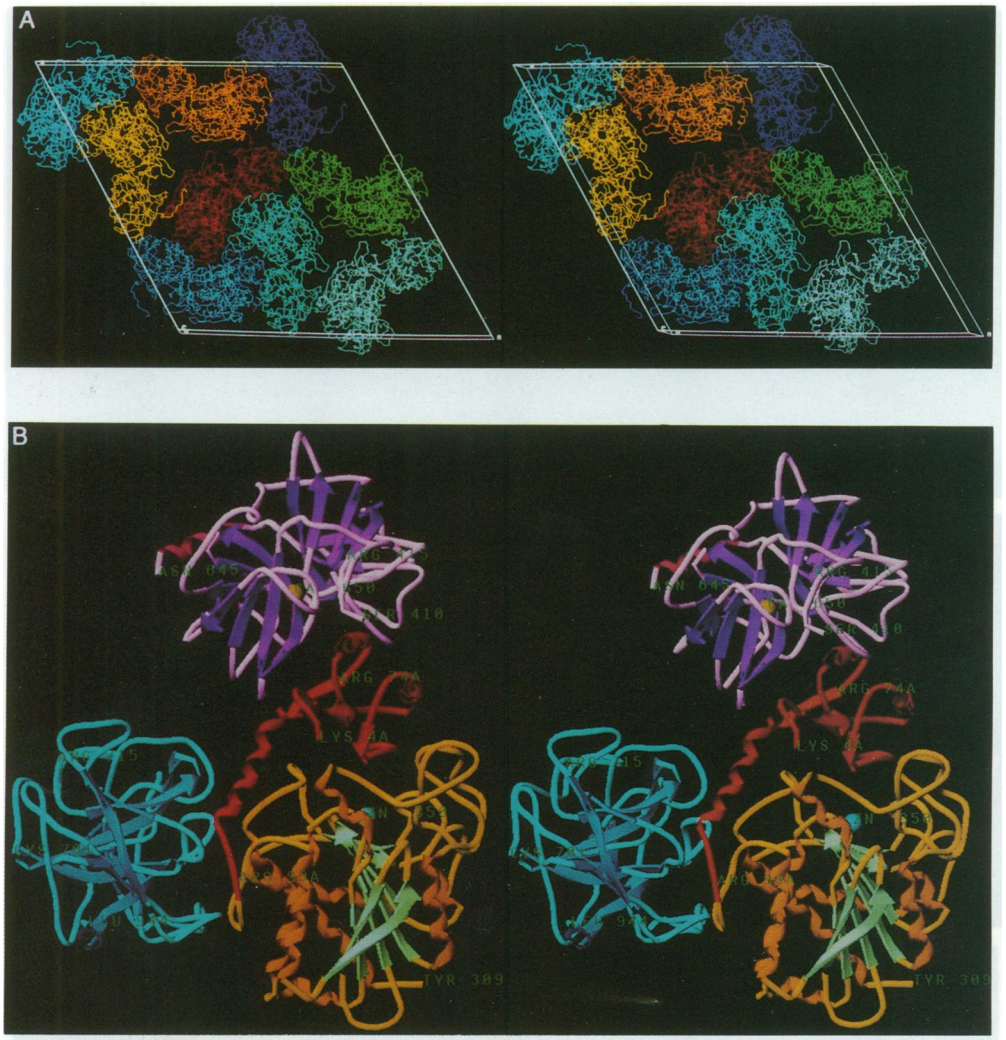

Fig. 1. Overall structure and intermolecular interactions within the TC. (**A**) Stereo picture displaying the packing of the nine TC oligomers present in one hexagonal unit cell. As can be noted, the freely extending N-terminal activation peptide of BPE (yellow oligotrimer) interacts with the CTGC moiety of a symmetry-related TC (red). (**B**) Stereo ribbon cartoon of the TC in the chosen standard orientation. The activation segment of PCPA (red) acts as a clamp for the CPA enzyme moiety (yellow), BPE (magenta), and CTGC (blue). The two localized ions, the catalytic zinc of CPA (blue sphere) and the structural calcium-ion of BPE (yellow sphere) are further displayed and labelled (ZN 350 and CA 650) in green, as are the N- and C-termini of each subunit and the main activation cleavage points [Arg99A and Arg74A of PCPA, Arg415 of BPE, and Arg715 of CTGC (tentatively traced)].

via His740, a polar hydrogen bond/salt-bridge network with Ser732 (a 'second triad', see Madison *et al.*, 1993). Probably, upon tryptic activation, the former aspartate residue becomes engaged in an internal salt-bridge with the liberated α-amino group of the new N-terminal residue Val716, shaping the active site, as happens in other homologous serine proteinases (Huber and Bode, 1978). The preceding segment from Gly884–Ser889 is completely disordered in CTGC (see Figure 2A), in contrast to CTGA, where the equivalent (but one residue shorter) segment is localizable.

Due to the glycine residue at position 916 and the bulky Thr926 residue, the specificity pocket of CTGC formed upon activation cleavage will have an extension between bovine chymotrypsin (with a glycine residue at the position equivalent to 926) and porcine pancreatic elastase (PPE; Meyer *et al.*, 1988), with a valine residue and a threonine residue at the positions equivalent to 916 and 926, respectively. This renders a pocket shaped to accommodate medium-sized residues, in agreement with bovine and porcine chymotrypsin-C preferences for leucine-P_1 residues (Folk and Schirmer, 1965).

The N-terminal segment of CTGC, clamped through Cys701 via Cys822 to the main molecular body, runs up to Pro708 along the molecular surface in an identical manner as in CTGA. The six residues (Ser711–Gly718, considering a two-residue deletion as compared with the equivalent CTGA sequence) flanking the scissile peptide bond Arg715–Val716 are, similarly to trypsinogen (Fehlhammer *et al.*, 1977), completely disordered in CTGC, making this activation cleavage site easily accessible for tryptic attack.

To the other side of the preformed CTGC pocket, the shifted 'autolysis loop' runs across its entrance, with Leu844 partially filling the void volume left. The following exposed part of this loop is very mobile and should be susceptible to (auto)catalytic digestion. The adjacent metal-lacking 'calcium-binding loop' exhibits a similar

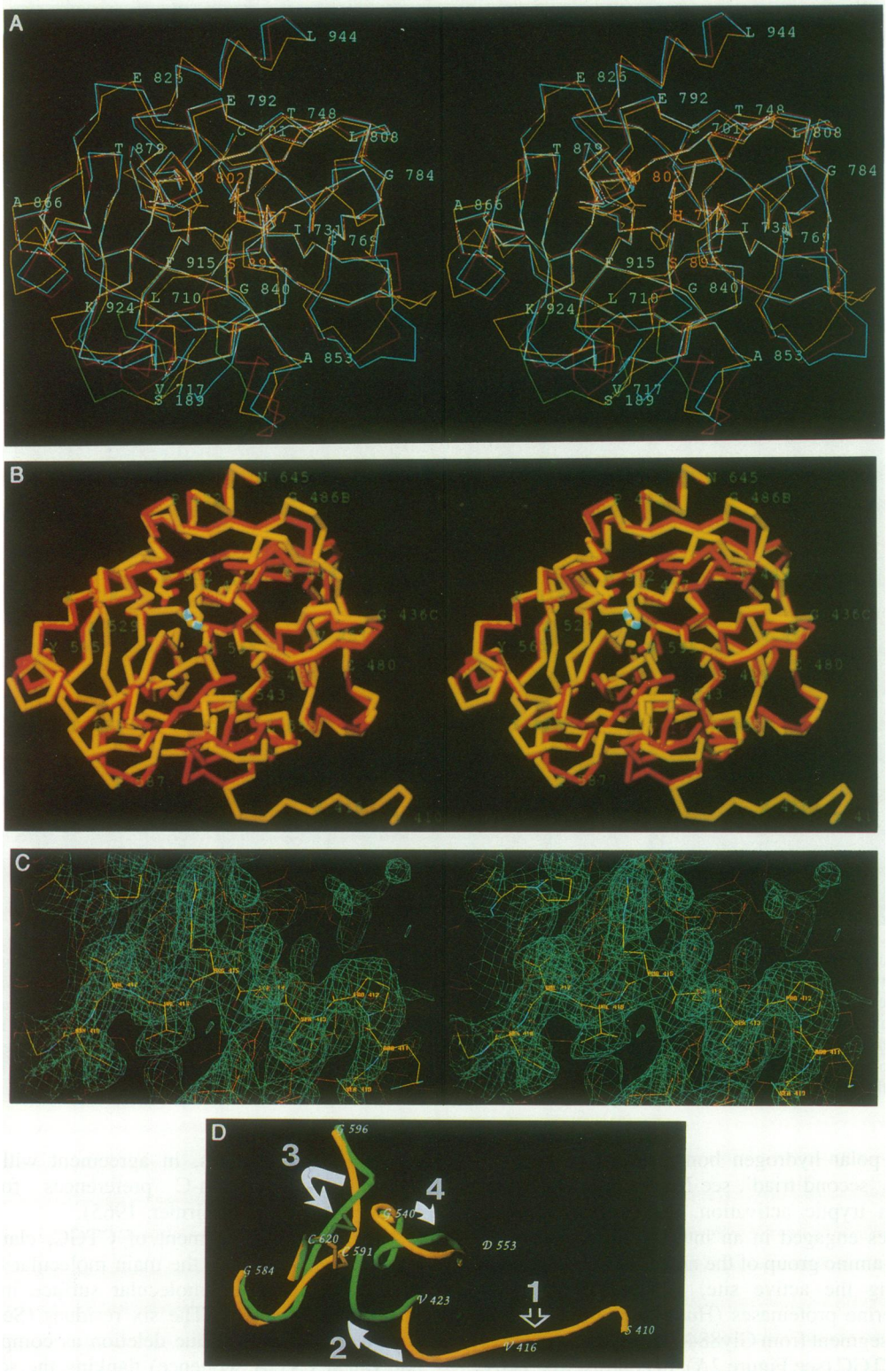

Fig. 2. The serine proproteinases within the TC. (**A**) Superimposition of the Cα-structures of CTGC (yellow; undefined segments in green), bovine α chymotrypsin A (PDB access code 5cha; sky blue), considered as a model for the active chymotrypsin C, and bovine CTGA (red; PDB access code 2cga) in the approximate standard orientation introduced by Fehlhammer and co-workers (1977) after optimal least-squares fitting. The side chains of the residues of the CTGC catalytic triad (Ser895, Asp802 and His757) are displayed and labelled, as are other residues of the CTGC structure. (**B**) Superimposition of the Cα-structures of BPE (yellow) and PPE (red), considered as a model for the active proteinase E, after optimal least-squares fitting. The side chains of the residues of the BPE catalytic triad (Ser595, Asp502 and His457) are displayed and labelled in green, as are other residues of the BPE structure and the position of Val16 in the mature enzyme (magenta). (**C**) Detail of the BPE structure around the N-terminal segment Ser410–Asn418 superimposed with the final $2F_o$–F_c electron density. This segment interacts with a symmetry-related CTGC molecule (displayed in the background in light brown). Same orientation as in (B). (**D**) Ribbon plot displaying the superimposition of the BPE (yellow) and PPE (green) structures just around the peptide segments involved in rearrangements upon tryptic activation. Some residues of the BPE structure are labelled, same orientation as in (C). Upon tryptic cleavage of bond Arg415–Val416 (1), the newly formed N-terminus is believed to project into the enzyme moiety (2). This conformational change may displace segment Gly584–Gly596 (3) and the 'autolysis-loop' Gly540–Asp553 (4) that can adopt the 'active' conformation.

conformation as found in trypsinogen (Fehlhammer *et al.*, 1977) and PPE (Meyer *et al.*, 1988) in spite of an extra residue.

According to most sequence and structure characteristics and also specificity, bovine CTGC can be identified as a hybrid between a typical (pro)elastase and a chymotrypsinogen A. It exhibits several insertions also present in PPE, and its sequence is significantly closer to porcine PPE (54% identity) or bovine BPE (56%) than to bovine CTGA (41%).

The structure of BPE subunit and comparison with PPE: an activation process hypothesis for the proteinase E family

Similar to CTGC, BPE exhibits the typical tertiary fold of chymotrypsin-like serine proteinases. According to structure and specificity of the mature enzyme, however, BPE is a typical proelastase (Kobayashi *et al.*, 1981b). Large conformational differences from the activated PPE (Meyer *et al.*, 1988) (no structure is available for a pancreatic proelastase) are only observed for three peptide segments arranged around the specificity pocket, mainly reflecting conformational changes occurring upon activation cleavage (see Figure 2B). Similar to CTGC and CTGA, segment Ser589–Asp594 (for the numbering convention of BPE residues, see Materials and methods) is folded into the molecule in BPE, and the coupled Asp594 side chain projects out of the molecule forming with Ser432 and His440 the 'second triad' (Madison *et al.*, 1993). The preceding segment Gly585–Gly590 is ordered in BPE, but takes a slightly different course compared with PPE, certainly a consequence of the large conformational change in the following segment (see Figure 2B). Disulfide bridge Cys591–Cys620 essentially allows for the considerable shift of Cys591 by a large internal reorganization (see Figure 2B).

'Below' this preformed pocket, the elongated N-terminal segment of BPE is arranged on the molecular surface (see Figures 2B–D) in a hairpin loop-like manner forming intramolecular contacts. From Ser410–Asn418, it runs simultaneously along the molecular surface of an adjacent symmetry-related CTGC molecule (see Figures 1A and 2C) between Arg415 and Asn418. Between Asn418 and Asp421, the BPE chain passes a regular 1,4 tight turn and then follows the equivalent chain segment of the activated PPE. Thus, the Gly419–Glu420 pair would seem to represent a hinge point in BPE, about which segment Val416–Gly420 would rotate after tryptic activation (see Figure 2D, steps 1 and 2) so that the newly liberated Val416 N-terminus could insert into the molecule and form the buried salt-bridge with Asp594; this salt-bridge formation is presumed to trigger the conformational changes (see Figure 2D, steps 3 and 4) leading to an open specificity pocket and a functional active site.

The 'autolysis loop' of BPE is fully ordered, with the chain segment following Gly542 taking a completely different course (up to Gly549) compared with activated enzymes such as PPE. The adjacent 'calcium-binding loop' is likewise well defined in BPE and exhibits a virtually identical conformation as in PPE (see Figure 2B), being the central calcium ion similarly heptahedrally coordinated.

Thus, with the exception of the three peptide segments affected by the reorganization upon activation cleavage, the BPE molecule exhibits a very similar structure to PPE. This close similarity is also reflected by a 54% sequence identity, compared with a still 41% identity towards bovine CTGA. The much higher 85% identity to human proproteinase E (Shen *et al.*, 1987) indicates, however, that this latter enzyme is the real equivalent of BPE in human beings, and that our BPE model can be considered a prototype for the human proenzyme. As with human proteinase E and also PPE, BPE should become a typical elastase after activation; Val616 and Thr626 (the same side chains are present at the equivalent positions of PPE) should likewise restrict the size of its S_1 pocket to small or medium-sized hydrophobic P_1 side chains (Kobayashi *et al.*, 1981a,b).

The BPE molecular structure in the TC differs strikingly from the recently published structure of the isolated bovine subunit III (Pignol *et al.*, 1994). Due to the absence of the first 13 residues (Phe405–Val417, according to our nomenclature), this truncated BPE variant was found to have a very low enzymatic activity (Wicker and Puigserver, 1981; Kerfelec *et al.*, 1986), conceivably due to the lack of the internal salt-bridge (equivalent to Val416–Asp594) required to stabilize an intact substrate-binding site and a correct active-site environment. The structural analysis of subunit III (Pignol *et al.*, 1994) revealed that the N-terminal segment is just defined and distinctly arranged from Pro424 onwards. However, the 'second triad' (Madison *et al.*, 1993) present in BPE as well as in the chymotrypsinogens and in trypsinogen (Fehlhammer *et al.*, 1977; Wang *et al.*, 1985) was not formed in subunit III; the 'calcium-binding loop', displaying no evidence for metal ion, had moved out of its 'normal' (i.e., in BPE, CTGA, trypsinogen and CTGC) position, occupying a surface region of the molecule corresponding to Val23–Pro24 in PPE; the adjacent 'autolysis loop' had turned towards the region 'normally' accommodating this calcium binding loop, and segment Gly515–Leu520 (according to our BPE nomenclature) was found with a different conformation. In comparison with PPE, these three surface loops display shifts in their α-carbon atoms up to almost 20 Å (see Pignol *et al.*, 1994). These great structural differences between BPE and subunit III seem mainly to depend on the absence of residues Arg415–Val417, involved in intramolecular contacts in BPE and apparently assisting in clamping the following segment up to Arg424 to the molecular surface. This bound N-terminal segment may directly stabilize the adjacent loops in their 'normal' position and not allow rearrangements as observed in subunit III. Additionally, lack of calcium in the crystallization solution of subunit III (Pignol *et al.*, 1994) could have destabilized the originally metal-binding 'calcium-binding loop', with further effects on the conformation of the surrounding loops.

The overall structural similarity of our BPE with CTGA, CTGC and trypsinogen shows, however, that the zymogens of the 'proteinase E family' do not substantially differ from the other related zymogens, and that the formulation of a new activation mechanism (Pignol *et al.*, 1994) is not justified.

Intra complex contacts and overall sequence of activation

In the TC, BPE interacts through 159 atom–atom contacts below 4 Å with PCPA, exclusively via the AS of the

Table I. Summary of data collection, processing and structure determination

Diffraction data								
No. of measured reflections	97896							
No. of unique reflections	30118							
R_{merge}[a]	0.063							
Completeness								
for data (20.3–2.57 Å)	86.7%							
for data (2.69–2.57 Å)	69.6%							
Patterson search and rigid-body refinement								
Cross rotation[b]	α	β	γ	Correlation function[d]				
Rotation function for								
PCPA	55.3	104.0	77.8	23.9				
	53.7	130.1	283.5	7.6				
BPE	99.8	53.5	158.9	14.4				
	103.4	79.5	341.5	11.5				
CTGC	91.3	52.1	1.2	10.8				
	0.2	66.8	346.3	8.4				
Translation[b,c]	α	β	γ	*x*	*y*	*z*	Correlation function[d]	Crystallographic *R* factor (%)
Independent translation function for								
PCPA	55.3	104.0	77.8	0.1564	0.1614	0.0000	36.3	45.4
	55.3	104.0	77.8	0.2605	0.4824	0.0000	21.2	50.3
BPE	99.8	53.5	158.9	0.0754	0.4538	0.0000	17.4	50.7
	99.8	53.5	158.9	0.0272	0.0713	0.0000	15.0	51.8
CTGC	91.3	52.1	1.2	0.5846	0.0452	0.0000	13.5	51.9
	91.3	52.1	1.2	0.9249	0.2763	0.0000	13.4	52.1
Dependent 3-body translation function[e]								
PCPA	55.3	104.0	77.8	0.1564	0.1614	0.0000	36.3	45.4
BPE	99.8	53.5	158.9	0.0784	0.4546	0.3529	46.4	41.9
CTGC	91.3	52.1	1.2	0.5908	0.0431	0.9622	47.4	42.1
Rigid-body refinement								
PCPA	56.4	104.6	77.3	0.1573	0.1605	0.0010		
BPE	101.4	55.5	158.4	0.0781	0.4547	0.3495	→ 54.6	39.0
CTGC	91.0	53.1	1.8	0.5902	0.0465	0.9679		

[a] $R_{merge} = \Sigma(\Sigma \mid I(h)_i - \langle I(h) \rangle \mid / \Sigma \langle I(h) \rangle)$; $I(h)_i$ is the observed intensity of the *i*-th measurement of reflection *h*, and $\langle I(h) \rangle$ the mean intensity of reflection *h*; calculated after loading, scaling and merging with rotavata/agrovata (CCP4, 1994).
[b] α, β, γ are given in Eulerian angles.
[c] *x,y,z* are given in fractional cell coordinates.
[d] Defined as in the AMoRe suite (Navaza, 1994).
[e] Correlation and *R* factor are cumulative after positioning of the first, second, and third molecules, respectively.
In the rotation and translation functions, the second highest peak is additionally displayed in each case for comparison.

latter. It essentially covers the second surface helix of the AS with its active-site groove, but also touches part of the N-terminal segment, the fourth (edge) strand and the N-terminal pole of the connecting helix. Only BPE residues of some exposed loop segments surrounding its active-site, i.e. loops which (except the autolysis loop) will presumably not undergo significant conformational changes upon tryptic activation as compared with the PPE structure, are involved. The same intermolecular contacts might occur in complexes formed with activated BPE, in agreement with experimental results showing that they are stable (F.X.Avilés and J.Vendrell, unpublished results). Steric hindrance around the second activation cleavage point by the adjacent BPE and stabilization through many intermolecular contacts will certainly help to prevent the susceptibility to tryptic attack of the AS of PCPA at this site, again in agreement with experimental results (Chapus *et al.*, 1987). In contrast, the activation cleavage site of BPE is located on the surface opposite to the PCPA contacting site (see Figure 1B) and is thus accessible to attacking trypsin molecules to the same extent as in related monomeric proelastases.

CTGC likewise covers the PCPA subunit with its active-site groove, but touches not only the AS, but also part of the enzyme domain of PCPA via some exposed loops under formation of 167 atom–atom contacts. It is noteworthy that CTGC is in particular clamped (via two hydrogen bonds) by means of Asn761 carboxamide to Arg99A N and O, contained in the primary activation scissile peptide bond Arg99A–Ala1 of PCPA. Thus, this principal PCPA

cleavage site seems not to be sterically accessible to tryptic cleavage in the TC, in agreement with earlier proposals (Michon *et al.*, 1991a). The activation cleavage site of the CTGC component is, as is that of BPE, directed away from the PCPA contacting surface and completely exposed to bulk solvent in the TC, and should be available for tryptic attack (see Figure 1B). Again, this would be in accordance with earlier findings, which seem to indicate the maintenance of the complex after CTGC activation (Uren and Neurath, 1972).

In conclusion, the resolution of the three-dimensional structure of the native TC helps the understanding of the conformational determinants in the proteolytic activation of the subunits. The AS of PCPA acts as a clamp for the two other serine proproteinases, which stabilize and sterically protect the first and the second activation cleavage sites of PCPA against tryptic attack. In contrast, the activation sites of both serine proteinase zymogens are exposed as in the monomeric state and are presumably susceptible to tryptic attack. Thus, these latter zymogens would be activated first (Kobayashi *et al.*, 1981b), which might slightly weaken or distort the TC and facilitate access of trypsin to the PCPA activation sites.

Materials and methods

Purification, crystallization, diffraction data collection and processing

Bovine native procarboxypeptidase A TC was purified as reported elsewhere by anionic exchange FPLC chromatography from pancreatic cetonic powder. The oligotrimer was crystallized using the sitting drop vapour diffusion method at 4°C. Drops containing 0.1 M $CaCl_2$, 0.05 M HEPES, and 15% (v/v) polyethylenglycol 400 (pH 7.1) as precipitating agent and protein (~15 mg/ml) were concentrated against reservoir buffer of the same precipitating agent. The crystals diffracting beyond 2.6 Å resolution belong to the space group R3 and have cell constants a = b = 188.5 Å, c = 82.5 Å (hexagonal setting). One TC oligomere is present per asymmetric unit. X-ray diffraction data were collected on a 180-mm MAR Research image plate detector attached to a Rigaku-Denki rotating anode CuK_α-generator operated at 5.4 kW at −12°C. The data were processed with MOSFLM v.5.23 (Leslie, 1987) and scaled, merged and reduced with rotavata/agrovata/truncate (CCP4, 1994) (see Table I).

Structure solution, sequence numbering conventions and representations

Orientation and position of the three molecules was found with the Patterson-search method (Huber, 1965). The starting models used were those of porcine PCPA (Guasch *et al.*, 1992) for the bovine homologous form, PPE (Meyer *et al.*, 1988) for BPE, and bovine CTGA (Wang *et al.*, 1985) for CTGC. The rotation and translation functions were computed with the AMoRe suite (Navaza, 1994) using data between 15 and 4 Å resolution (see Table I). Rigid body refinement of the properly oriented and positioned molecules was performed with AMoRe, whereby the crystallographic R factor (defined as $\Sigma\{|F_{obs}|-|F_{calc}|\}/\{\Sigma|F_{obs}|\}$) dropped from 0.421 to 0.390. The model was inspected and corrected against ($2F_{obs}$–F_{calc}) and (F_{obs}–F_{calc}) Fourier maps in successive cycles comprising manual model building (using TURBO-Frodo version 5.0a, Bio-Graphics, Marseille, France) at a Silicon Graphics workstation and crystallographic refinement using X-PLOR (Brünger *et al.*, 1987). The crystallographic R factor for the final model as refined against 29060 reflections between 8.0 and 2.6 Å resolution is 0.198 (free R factor is 0.297); the r.m.s deviations from target values for bond lengths and angles are 0.013 Å and 1.895° respectively. PCPA residues have been numbered according to PDB access code 1pca (Guasch *et al.*, 1992). For rationality reasons, in the BPE residue numbering 400 counts have been added to the equivalent CTGA numbering, that is, Ser195, Asp102 and His57 of CTGA are equivalent to Ser595, Asp502, and His457 in BPE. The CTGC numbering starts with residue Cys701 (700 counts added to the corresponding CTGA residues; the CTGC catalytic triad therefore consists of Ser895, Asp802, and His757). Segments Ser711–Gly718, Tyr846–Asn848 and Gly885–Ser889 of CTGC are not defined by appropriate electron density; these 16 residues and seven surface-located side chains have nevertheless been tentatively traced, in order to preserve chain continuity, and set to zero occupancy. The first five residues of BPE (Phe405–Phe409; the BPE N-terminal activation peptide is shorter by four residues than that of CTGA) are not defined by electron density and have not been traced. Therefore, the final model comprises the residues from PCPA (403 residues, labelled 4A–99A and 1–309), BPE (248 residues out of 253, labelled 410–645 considering insertions/deletions), CTGC (251 residues, labelled 701–945), one zinc and one calcium ion (labelled 350 and 650, respectively), and 359 solvent molecules (labelled W100–W458). An examination of main-chain dihedral angles (performed with PROCHECK v. 3.0, implemented in CCP4 Program Suite release 2.10; CCP4, 1994) showed three residues in disallowed regions. These residues, Glu720, Ala852 and Ala890, are part of flexible regions, just following the undefined segments mentioned above. The average temperature factor for the 7239 active non-hydrogen protein atoms is 36.6 $Å^2$.

Structural superimpositions were performed with LSQKAB, implemented in the CCP4-package, and TURBO-Frodo, and displayed with the latter program on a Silicon Graphics workstation. All pictures have been made with TURBO-Frodo. The final coordinates will be deposited with the Brookhaven Protein Data Bank and will be released one year after publication (PDB access code 1pyt).

Acknowledgements

The kind help at early stages of protein purification and structure analysis provided by S.Ventura, V.Villegas, J.Vendrell, J.Navaza, H.Brandstetter and I.Fita, as well as the financial help provided by fellowships ERBCHBICT920053 (Human Capital and Mobility Programme of the European Union) and ALTF 371-1992 (EMBO long-term fellowship) is acknowledged. This work has been further supported by grants BIO92-0458 (CICYT, Ministerio de Educación, Spain) and ERBCHRXC-T940535 (Human Capital and Mobility Network of the European Union), and by the Centre de Referència de Biotecnologia (Generalitat de Catalunya, Spain). Thanks, too, to A.Puigserver for providing the unpublished chymotrypsinogen C sequence and to T.Mather for helpful discussions.

References

Avilés,F.X., Vendrell,J., Guasch,A., Coll,M. and Huber,R. (1993) Advances in metallo-procarboxypeptidases. Emerging details on the inhibition mechanism and on the activation process. *Eur. J. Biochem.*, **211**, 381–389.

Bode,W. and Huber,R. (1986) Crystal structures of pancreatic serine endopeptidases. In Desnuelle,P., Sjöström,H. and Norèn,O. (eds), *Molecular and Cellular Basis of Digestion.* Elsevier, Amsterdam, pp. 213–234.

Brown,J.R., Cox,D.J., Greenshields,R.N., Walsh,H., Yamasaki,M. and Neurath,H. (1961) The chemical structure and enzymatic functions of bovine procarboxypeptidase A. *Proc. Natl Acad. Sci. USA*, **47**, 1554–1562.

Brünger,A.T., Kuriyan,K. and Karplus,M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science*, **235**, 458–460.

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Chapus,C., Kerfelec,B., Foglizzo,E. and Bonicel,J. (1987) Further studies on the activation of bovine pancreatic procarboxypeptidase A by trypsin. *Eur. J. Biochem.*, **166**, 379–385.

Coll,M., Guasch,A., Avilés,F.X. and Huber,R. (1991) Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. *EMBO J.*, **10**, 1–9.

Fehlhammer,H., Bode,W. and Huber,R. (1977) Crystal structure of bovine trypsinogen at 1.8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J. Mol. Biol.*, **111**, 415–438.

Folk,J.E. and Schirmer,E.W. (1965) Chymotrypsin C. I. Isolation of the zymogen and the active enzyme: Preliminary structure and specificity studies. *J. Biol. Chem.*, **240**, 181–192.

Freisheim,J.H., Walsh,K.A. and Neurath,H. (1967) The activation of bovine procarboxypeptidase A. II. Mechanism of activation of the succinylated enzyme precursor. *Biochemistry*, **10**, 3020–3028.

Guasch,A., Coll,M., Avilés,F.X. and Huber,R. (1992) Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison

of the A and B zymogens and their determinants for inhibition and activation. *J. Mol. Biol.*, **224**, 141–157.
Hartley,B.S. (1964) Amino-acid sequence of bovine chymotrypsinogen A. *Nature*, **201**, 1284–1287.
Huber,R. (1965) Die automatisierte Faltmolekülmethode. *Acta Crystallogr.*, **A19**, 353–356.
Huber,R. and Bode,W. (1978) Structural basis of the activation and action of trypsin. *Accts. Chem. Res.*, **11**, 114–122.
Kerfelec,B., Cambillau,C., Puigserver,A. and Chapus,C. (1986) The inactive subunit of ruminant procarboxypeptidase A-S6 complexes. Structural basis of inactivity and physiological role. *Eur. J. Biochem.*, **157**, 531–538.
Kobayashi,R., Kobayashi,Y. and Hirs,C.H.W (1981a) The specificity of porcine pancreatic protease E. *J. Biol. Chem.*, **256**, 2460–2465.
Kobayashi,Y., Kobayashi,R. and Hirs,C.H.W (1981b) Identification of zymogen E in a complex with bovine procarboxypeptidase A. *J. Biol. Chem.*, **256**, 2466–2470.
Leslie,A.G.W. (1987) Profile fitting. In Helliwell,J.H., Machin,P.A. and Papiz,M.Z. (eds), *Computational Aspects of Protein Crystal Data Analysis*. Proceedings of the Daresbury Study Weekend, 23–24 January.
Madison,E.L., Kobe,A., Gething,M.-J., Sambrook,J.F. and Goldsmith,E.J. (1993) Converting tissue plasminogen activator to a zymogen: a regulatory triad of Asp-His-Ser. *Science*, **262**, 419–421.
Meyer,E., Cole,G., Radahakrishnan,R. and Epp,O. (1988) Structure of native porcine pancreatic elastase at 1.65 Å resolution. *Acta Crystallogr.*, **B44**, 26–38.
Michon,M., Sari,J.C., Granon,S., Kerfelec,B. and Chapus,C. (1991b) Microcalorimetric investigation of the interactions between the subunits of the bovine pancreatic procarboxypeptidase A complex. *Eur. J. Biochem.*, **201**, 217–222.
Michon,T., Granon,S., Sauve,P. and Chapus,C. (1991a) The activation peptide of pancreatic procarboxypeptidase A is the keystone of the bovine procarboxypeptidase A-S6 ternary complex. *Biochem. Biophys. Res. Commun.*, **181**, 449–455.
Navaza,J. (1994) AMoRe: An automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.
Neurath,H. (1989) Proteolytic processing and physiological regulation. *Trends Biochem. Sci.*, **14**, 268–271.
Pascual,R., Vendrell,J., Avilés,F.X., Bonicel,J., Wicker,C. and Puigserver,A. (1990) Autolysis of proproteinase E in bovine procarboxypeptidase A ternary complex gives rise to subunit III. *FEBS Lett.*, **277**, 37–41.
Pignol,D., Gaboriaud,C., Michon,T., Kerfelec,B., Chapus,C. and Fontecilla-Camps, J.C. (1994) Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a highly structured zymogen E. *EMBO J.*, **8**, 1763–1771.
Puigserver,A. and Desnuelle,P. (1977) Reconstitution of bovine procarboxypeptidase A-S6 from the free subunits. *Biochemistry*, **16**, 2497–2501.
Puigserver,A., Chapus,C. and Kerfelec,B. (1986) Pancreatic exopeptidases. In Desnuelle,P., Sjöström,H. and Norèn,O. (eds), *Molecular and Cellular Basis of Digestion*. Elsevier, Amsterdam, pp. 235–247.
Shen,W., Fletcher,T.S. and Largman,C. (1987) Primary structure of the human pancreatic protease E determined by sequence analysis of the cloned mRNA. *Biochemistry*, **26**, 3447–3452.
Uren,J.R. and Neurath,H. (1972) Mechanism of activation of bovine procarboxypeptidase A S5. Alterations in primary and quaternary structure. *Biochemistry*, **11**, 4483–4492.
Vendrell,J., Cuchillo,C.M. and Avilés,F.X. (1990) The tryptic activation pathway of monomeric procarboxypeptidase A. *J. Biol. Chem.*, **265**, 6949–6953.
Venot,N., Sciaky,M., Puigserver,A., Desnuelle,P. and Laurent,G. (1986) Amino acid sequence and disulfide bridges of subunit III, a defective endopeptidase present in the bovine 6S procarboxypeptidase A complex. *Eur. J. Biochem.*, **157**, 91–99.
Wang,D., Bode,W. and Huber,R. (1985) Bovine chymotrypsinogen A. X-ray crystal structure analysis and refinement of a new crystal form at 1.8 Å resolution. *J. Mol. Biol.*, **185**, 595–624.
Wicker,C. and Puigserver,A. (1981) Further studies on subunit III of bovine pancreatic procarboxypeptidase A. *FEBS Lett.*, **128**, 13–16.

Received on May 24, 1995; revised on June 23, 1995